Response to comments on "PMIF v1.0: an inversion system to estimate the potential of satellite observations to monitor fossil fuel CO2 emissions" by Y. Wang et al.

We thank the referee for reviewing our manuscript and for his valuable comments and suggestions. Please find attached a point-by point reply (in black) to each of the comments raised by the referee (in blue) with legible text and figures organized along the text. For your convenience, changes in the revised manuscript are highlighted with dark red. All the pages and line numbers correspond to the original version of text.

1 Overview:

Review of "PMIF v1.0: an inversion system to estimate the potential of satellite observations to monitor fossil fuel CO2 emissions" by Wang et al. Wang et al. present an OSSE framework to estimate error reductions for a proposed satellite. It's based on a Gaussian plume that they run for many emission hotspots. They've done this over a large domain (globally) at fairly high spatial resolution (2 km). The work is interesting but the description of the methods could use quite a bit of work. There are some important steps in the actual implementation that are quite convoluted. Fixing this seems like a critical for publication in a journal focused on geoscientific model development. I suggest major revisions for the manuscript.

Response:

We carefully revised our manuscript following the comments and suggestions. We think that the revised manuscript explained the steps of the method much clearer.

2 Comments:

2.1 Solution to their inversion

I'd prefer the authors not use A as the posterior covariance matrix, I usually think of A as the averaging kernel. This is particularly confusing because you are solving for emission reductions that are the diagonals of the averaging kernel matrix.

Response:

We are aware of that $\mathbf{A}$ (or AK) is used for averaging kernel in the community of satellite retrievals (Boesch et al., 2011; Cogan et al., 2012; O'Dell et al., 2012; Wu et al., 2018b; Yoshida et al., 2011). We also know that in some books on atmospheric inversion, $\mathbf{A}$ is used to represent "the sensitivity of the optimal estimate to the true state" and is also called averaging kernel (such as in Daniel Jacob's "Lecture on Inverse modelling" http://acmg.seas.harvard.edu/education/jacob_lectures_inverse_modeling.pdf), where $\mathbf{S}_A$ and $\hat{\mathbf{S}}$ are used for prior and posterior uncertainty. The $\mathbf{P}_f$ (with f for forecast) and $\mathbf{P}_a$ (with a for analysis) notations from the weather data assimilation community are also sometimes used in the GHG flux inverse modeling community. But flux inversion does not involve forecast steps, so $\mathbf{A}$ is widely used to represent the posterior covariance matrix in massive studies on atmospheric inversion (Broquet et al., 2018; Chevallier et al., 2005; Rayner et al., 2019) and in Peylin et al. (2013) which synthetizes the contributions from a wide range of inverse modeling groups. In particular, the review on atmospheric inversions by Rayner et al. (2019) tries to build consensus in the inverse modeling community regarding the notation and encourages the use of $\mathbf{A}$ for posterior uncertainty covariance matrices. So in this study, we choose to follow this

suggestion.

In addition, we want to clarify here we are not only solving for the diagonals of the posterior uncertainty matrix **A**. When we investigate the posterior uncertainty at daily and annual scales (Sect. 3.3 and 3.4), we account for the temporal auto-correlations in the prior uncertainty matrix **B**, which are the off-diagonals. The resulting **A** matrix is not a diagonal matrix, and we aggregate the **A** matrix at the scales of 3 h and 21 h time windows to daily and annual scales accounting for the off-diagonal entries of **A**.

In any case, Supplemental Section 1 presents what the authors are actually doing, which differs from the equations they present in Eq. 1 and 2. In Supplemental Section 1 the authors present a derivation that is both important and convoluted. It's unclear if this is something the authors devised themselves or if it follows from other work. Typically when people decompose error covariance matrices into spatial and temporal components they use a Kronecker product (e.g., Yadav & Michalak, GMD 2013). The Kronecker product greatly reduces the computational expense. The assumptions that go into a Kronecker product are also easy to follow because it is widely used. It's also amenable to sparse matrices (I'm assuming the authors are using sparse matrices). I think the authors should remove Equation 2 and bring Supplemental Section 1 into the main text. Supplemental Section 1 is important because this is what they are actually doing. This seems like the main contribution to me.

Response:

Eq. 1 and 2 explains the primary principle of atmospheric inversion and PMIF. We explained in Ln 143 "PMIF is an analytical inversion system that solves for Eq. (1) by building the different matrices involved in this equation." In the revised manuscript, we update this sentence with "PMIF is an analytical inversion system that solves for Eq. (1) or for an approximation of this equation (when accounting for temporal correlations in **B**) by building the different matrices involved in this equation.".

PMIF attempts at solving for Eq. 1 as well as possible. The PMIF-Paris OSSE and the experiment Exp-NoCor in PMIF-Globe solve for Eq. 1. Accounting for the temporal correlation in prior uncertainties (**B**) in other experiments in PMIF-Globe prevents from applying Eq. 1, and the Supplemental Section 1 (in the revised manuscript, it will be moved to the main text) explained how an approximation of the full **A** is derived in practice. We regularly use the Kronecker product for modeling spatio-temporal correlations, or temporal correlations at different temporal scales in inversions, e.g. in Wang et al. (2018), or to reduce the size of **B** matrices to be inverted in variational inversions, e.g. Broquet et al. (2011). But the Kronecker product cannot help to solve for the inversions of the $\mathbf{B}^{-1}+\mathbf{M}^T\mathbf{R}^{-1}\mathbf{M}$ matrix whose dimension is huge (on the order of of $10^7 \times 10^7$ since the control vector consist of $365 \times 2 \times 11,314 = 8.3 \times 10^6$) and whose non-diagonal terms can expand far from the diagonal when accounting for temporal correlations in PMIF-Globe inversions. In addition, due to the large number of satellite observations, $\mathbf{MBM}^T+\mathbf{R}$ is even larger, being 2.7×10^7 by 2.7×10^7. In Yadav and Michalak (2013), they computed and inverted the full $\mathbf{MBM}^T+\mathbf{R}$ matrix despite using the Kronecker product to gain computational efficiency for other diagnostics. But computing and inverting $\mathbf{B}^{-1}+\mathbf{M}^T\mathbf{R}^{-1}\mathbf{M}$ or $\mathbf{MBM}^T+\mathbf{R}$ in PMIF would require approximately 6000 TB of RAM, which is too much for the super computers in our lab.

In addition, as explained above, we need to aggregate the posterior uncertainty matrix **A**

at daily and annual scales. At these scales, **A** integrates the constraints from the temporal correlations in **B** and the spatial overlapping of plumes $\mathbf{M^T R^{-1} M}$, and the spatial overlapping of plumes differs from day to day depending on the wind fields. We do not see that $\mathbf{B^{-1} + M^T R^{-1} M}$ or **A** is necessarily a sparse matrix that can be computed with the Kronecker product.

Therefore, we actually devised the algorithm in Supplemental Section 1 by ourselves to approximate the diagonal of the full **A**. We admit that this method does not solve for **A** exactly, but only approximates the **A** at the scales we are interested in.

To prove that this approximation is good, we conduct an experiment with the ASS configuration of prior uncertainty where the inversion period and domain are limited to 6 months and to the Benelux, a region with high emission density and in which the 95 emission clumps are close to each other (Fig. R1a). It is reasonable to assume that if the approximation of the posterior uncertainty of emissions from clumps within this region (because we ignore the filtering of information from different spatial overlaps of plumes on different days, see the method) is good, clumps outside this inversion domain will have very marginal impact on the results for the clumps in Benelux. In this case, the full **A** can be explicitly derived based on Eq. 1. We compare this exact computation of the full **A** (Inv-fullA) to that obtained with the approach we proposed (Inv-2step). Figure R1b shows the posterior uncertainties in the emission budgets over individual time windows 8:30-11:30 for an exemplary clump (Antwerp) from the two computations. The results from the two computations are very close, except for very few days, and the aggregated uncertainty in emission budget for the whole period differ by less than 0.1%. This confirms that our method provides a good approximation of **A** at daily to annual scales for individual clumps with reasonable accuracy.
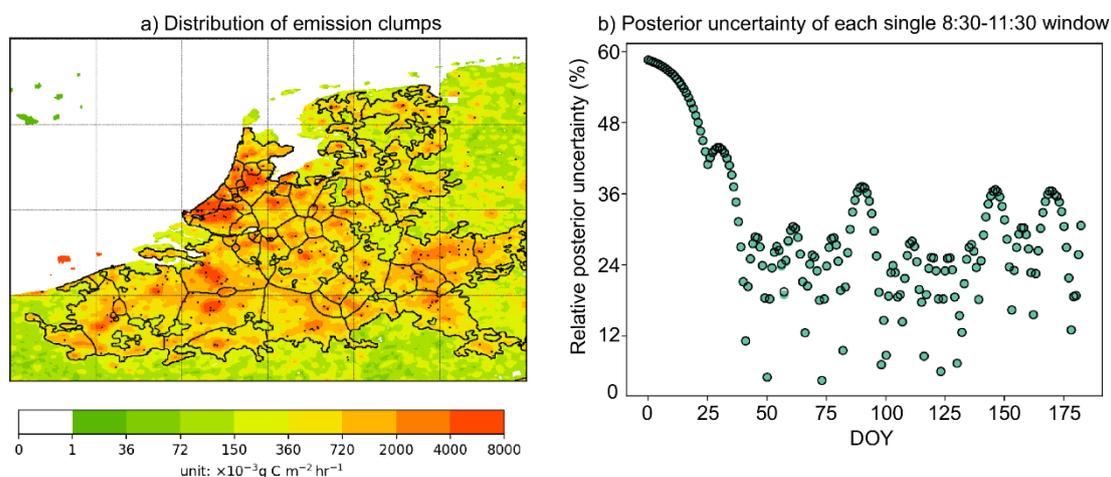

Figure R1 a) Distribution of emission clumps in the Benelux region that we account for in the Inv-fullA and Inv-2step inversions. The solid lines depict the boundaries of clumps. b) Posterior uncertainty of each single 8:30-11:30 window for Antwerp clump during the first half of the year. The green dots are the results from Inv-fullA, and the circles are the results from Inv-2step.

To address the reviewer's concern, we revised the manuscript by moving Supplemental Section 1 to the main text and slightly improving it:

"In this second set of OSSEs, PMIF-Globe, we conduct inversions for all the clumps over one year. However, the large sizes of the control vector, of the observation vector and of the associated covariance matrices prevent the derivation of a full **A** for all the clumps and all the

time windows using Eq. (1). In PMIF, we thus propose and apply a two-step computation that approximates Eq. (1). This computation assumes that the system has a limited capability to improve the separation between plumes from distinct clumps on a given day by crossing the information obtained from different days. In that sense, the inversion considers the uncertainty reduction obtained for individual days when considering all the clumps together (first step, see below) before focusing on individual clumps to account for temporal correlations in the prior uncertainty (the second step, see below). In other words, we assume that when crossing information between different time windows for a given clump, the impact of filtering information from different spatial overlaps of plumes on different days is relatively smaller than that of temporal auto-correlation in the prior uncertainty. It is proven that this method provides a good approximation of **A** at daily to annual scales for individual clumps (Supplementary text S1).

In the first step, Eq. (1) is applied to each $10°\times10°$ spatial inversion windows on each day separately (corresponding to an 8:30-11:30 time window for clumps within the spatial inversion windows), by using the corresponding blocks in **B**:

$$\mathbf{A}_{\text{spt},i,j} = \left(\mathbf{B}_{\text{spt},i,j}^{-1} + \mathbf{M}_{\text{spt},i,j}^{\mathrm{T}}\mathbf{R}_{\text{spt},i,j}^{-1}\mathbf{M}_{\text{spt},i,j}\right)^{-1} \tag{6}$$

Where i is the ith spatial inversion window and j is the jth day during one year. Here, $\mathbf{B}_{\text{spt},i,j}$ is a diagonal matrix that only contains the variances of prior uncertainties in emissions during 8:30-11:30 for the clumps within the inversion window. $\mathbf{M}_{\text{spt},i,j}$ accounts for the spatial overlap of plumes generated from nearby clumps. Then we derive a "instant" $\mathbf{M}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{M}$ (denoted as $\mathbf{M}_{\mathrm{I},j,k}^{\mathrm{T}}\widehat{\mathbf{R}_{\mathrm{I},j,k}^{-1}}\mathbf{M}_{\mathrm{I},j,k}$) for a given clump k at each 8:30-11:30 time window:

$$\mathbf{M}_{\mathrm{I},j,k}^{\mathrm{T}}\widehat{\mathbf{R}_{\mathrm{I},j,k}^{-1}}\mathbf{M}_{\mathrm{I},j,k} = \left(a_{\text{spt},i,j}(k)^{-1} - b_{\text{spt},i,j}(k)^{-1}\right)^{-1} \tag{7}$$

Where $a_{\text{spt},i,j}(k)$ is a scalar from $\mathbf{A}_{\text{spt},i,j}$ representing the variance of posterior uncertainty of emission from clump k in ith spatial inversion window and in 8:30-11:30 time window on day j obtained by Eq. (6), and $b_{\text{spt},i,j}(k)$ is the scalar from $\mathbf{B}_{\text{spt},i,j}$ representing the variance of prior uncertainty for the same control variable.

In the second step, the inversion is conducted for each clump k separately, considering the correlation in time in **B**, using $\mathbf{M}_{\mathrm{I},j,k}^{\mathrm{T}}\widehat{\mathbf{R}_{\mathrm{I},j,k}^{-1}}\mathbf{M}_{\mathrm{I},j,k}$ derived from the first step:

$$\mathbf{A}_{tmp,k} = \left(\mathbf{B}_{tmp,k}^{-1} + \begin{bmatrix} \mathbf{M}_{\mathrm{I},1,k}^{\mathrm{T}}\widehat{\mathbf{R}_{\mathrm{I},1,k}^{-1}}\mathbf{M}_{\mathrm{I},1,k} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{M}_{\mathrm{I},n,k}^{\mathrm{T}}\widehat{\mathbf{R}_{\mathrm{I},n,k}^{-1}}\mathbf{M}_{\mathrm{I},n,k} \end{bmatrix}\right)^{-1} \tag{8}$$

Where n=366×2, representing the time windows for 8:30-11:30 and for the rest 21 hours on the 366 days of one year (2008). $\mathbf{B}_{tmp,k}$ is the covariance matrix accounting for the temporal auto-correlation in the prior uncertainty for a single clump:

$$\mathbf{B}_{tmp,k} = \begin{bmatrix} \sigma_{t1}^2 & cov(\varepsilon_{t1},\varepsilon_{t2}) & \dots & cov(\varepsilon_{t1},\varepsilon_{tn}) \\ cov(\varepsilon_{t1},\varepsilon_{t2}) & \sigma_{t2}^2 & \dots & cov(\varepsilon_{t2},\varepsilon_{tn}) \\ \vdots & \vdots & \ddots & \vdots \\ cov(\varepsilon_{t1},\varepsilon_{tn}) & cov(\varepsilon_{t2},\varepsilon_{tn}) & \dots & \sigma_{tn}^2 \end{bmatrix} \tag{9}$$

In PMIF-Globe, we first conduct the inversion in which the prior uncertainty has no

temporal auto-correlation (Exp-NoCor)…*"*

<u>Response:</u>

We agree that "pseudo" can be misleading. In the revised manuscript, we bring the Supplemental Section 1 in Sect. 2.7.2 and replace "pseudo $\mathbf{M}^T\mathbf{R}^{-1}\mathbf{M}$" with "instant $\mathbf{M}^T\mathbf{R}^{-1}\mathbf{M}$".

The authors should change the title. It's not an inversion framework as they are not estimating fluxes.

<u>Response:</u>

As it stands, this tool can be used to process individual samples of pseudo prior fluxes and pseudo observations and compute pseudo posterior fluxes to assess error reductions to a pseudo truth. All the numerical objects needed to apply Eq.2 are built in this system as reflected by its description. However, if the errors injected in such OSSEs with explicit pseudo data are consistent with the statistics of uncertainties know by the inversion system, the statistics of errors in the flux estimates are fully characterized by $\mathbf{A}$ (since the observation operator is linear), whose direct computation is thus the best index of the potential of the inversion and of a given observation network (Wang et al., 2018). This is why we focus on such a computation here. This computation of $\mathbf{A}$ is actually a standard atmospheric inversion computation. Technically speaking, the PMIF can also be used to assimilate real data to produce estimate of the fluxes. Based on such considerations, the PMIF is an atmospheric inversion system like others so we wish to keep this label for clarity.

Of note is that this tool is mainly designed for OSSEs and would require some adaptations and extensions to process real satellite images over the period of data availability, to remove the XCO_2 background concentrations underlying the detected plumes, and maybe to more properly cope with errors in the modeling of the plumes (see our answer to the next comment) than just inflating the $\mathbf{R}$ matrix. However, such a limited account for model error in real experimental conditions is a traditional weakness of atmospheric inversion systems and other systems mainly designed for OSSEs that have always been named atmospheric inversion systems (Rayner et al., 2014; Wu et al., 2016).

2.2 Justification on the use of a Gaussian plume
Real plumes are only Gaussian in the time-averaged sense. The satellite observations provide a snapshot in time that likely would not be Gaussian. I think the authors need to provide some justification as to why a Gaussian plume is appropriate for data that is not time-averaged. A Gaussian plume may give a reasonable upper bound on the uncertainty reduction, but will likely induce systematic biases if implemented operationally. These potential biases should be discussed.

<u>Response:</u>

We agree with the reviewer that instant image of real plumes do not always follow a Gaussian shape: because of the turbulence close to the source, of the 3D variations in the wind

field, especially over the long distances, and of multiple other parameters (like variations in the topography, the complexity of vertical mixing etc.). However:

1) we stress, again, that the PMIF was not designed to process real data, but for OSSEs. The primary driver of the scores of posterior uncertainties and of uncertainty reduction in the PMIF which are the target of the OSSEs and of this system is the spatial extent and amplitude of the plumes, and the parameters of the Gaussian model in the PMIF are such that they fairly reproduce those from more complex models. This had been checked based on our comparisons between the results from the PMIF-Paris and from Broquet et al. (2018) as explained in the Section 3.1 and in the supplementary material.

2) the complex variations in real plumes that make them hardly Gaussian also hamper their modeling with complex mesoscale atmospheric transport models; this explains why many of the recent inversions of CO_2/CH_4 plant and city emissions that have been conducted based on OCO-2/TROPOMI data use Gaussian models or a Gaussian approximation of the shape of the plume to apply direct flux computations in the data (Nassar et al., 2017; Reuter et al., 2019; Zheng et al., 2020).

3) The study by Prunet et al. (2020) (the talk available at https://cdn.eventsforce.net/files/ef-xnn67yq56ylu/website/9/5_734__pascal_prunet-_plume_detection_and_characterization_from_xco__imagery-evaluation_of_gaussian_methods_for_quantifying_plant_and_city_fluxes.pptx) even indicates that Gaussian models fit the plumes from "true" mesoscale models well enough (so that the inversions using the Gaussian model can provide a good estimate of the emissions) for a good part of the typical atmospheric conditions encountered around the set of European cities and plants they investigated.

The choice of the Gaussian plume model in the PMIF was definitely linked to its light computation cost while using 2 km resolution observations and solving for emissions at a high resolution across the globe and a year. We think this choice does not bias the results given the different considerations listed above.

To better address this discussion about the Gaussian plume model in the manuscript, we revised it by:

1) revising Ln 101: "Therefore, in this study, we develop a Plume Monitoring Inversion Framework (PMIF) and conduct a set of Observing System Simulation Experiments (OSSEs) to assess, for the first time, the performance of a satellite instrument to monitor the emissions of all the clumps across the globe and over a whole year. The imager studied has the foreseen characteristics of the individual satellites of the forthcoming CO2M mission. It would be a high-resolution spectrometer, with 2 km × 2 km resolution pixels and a swath of 300 km, and it would be placed on a sun-synchronous orbit ensuring global coverage in 4 days. The PMIF inversion system relies on the list of clumps extracted by Wang et al. (2019) from the ODIAC inventory, on a Gaussian plume model to simulate the XCO_2 plumes generated by the emissions from these clumps, on an analytical inverse modeling framework, and on a combination of overlapping assimilation windows to solve for the inversion problem over the globe and a full year. It also addresses the question of temporal extrapolation that is needed to generate estimates of annual emissions from the information of a limited number of time windows for which emissions are well constrained by the direct satellite images, by accounting for the temporal auto-correlation of the prior uncertainties. The performance is assessed in terms of the

uncertainties in the emissions (Sect. 2.1) at different scales. The PMIF uses a Gaussian plume model at the local scale to ensure that the computation cost is affordable. Such a model can often hardly fit with actual plumes over the distances considered in this study (due to variations in the wind field, topography, vertical mixing etc. over such distances) but is shown, when driven with suitable parameters, to provide a satisfactory simulation of the plume extent and amplitudes, which appear to be the main drivers of the targeted computations of uncertainties in the emissions in our OSSE framework (as shown in section 3.1). In PMIF, we also ignore the impact of some sources of uncertainties on the inversion of emissions, including systematic errors on the XCO_2 retrievals, the impact of uncertainties in diffuse anthropogenic emissions outside clumps, in non-fossil CO_2 fluxes (within and outside clumps), and in the spatial and temporal variations of emissions within the clump and the short time windows that the inversion aims to solve. These impacts are discussed in detail afterwards."

2) revising Ln 148-157: "We use a Gaussian plume model (Sect. 2.4) to simulate the atmospheric transport at a spatial resolution consistent with that of the XCO_2 measurements from the planned CO_2 imager and with the highly heterogeneous distribution of emissions. Compared with complex 3-D atmospheric transport models, Gaussian plume models have a very low computational cost, making the global assessment of posterior uncertainty and uncertainty reduction at the scale of emissions clumps from the assimilation of high resolution data feasible. However, since a Gaussian plume model provides a highly simplified approximation of the atmospheric transport from emission clumps, we need to verify that its use in the PMIF yields estimates of the uncertainties in the inverted emissions that are consistent with those that would be based on more complex models. Therefore, we first compare the results for Paris from PMIF against those acquired based on a 3-D Eulerian mesoscale atmospheric transport model by Broquet et al. (2018)… "

The authors should give more explanation of σj . There are two parameters in a Gaussian plume model and they spend one line talking about σj : "The σj is a function of downwind distance i and atmospheric stability parameter. We take the form for σj from Ars et al. (2017).".

Response:

To clarify our set-up of the parameters in the Gaussian plume model used here, we revise the sentence in Ln 225: "The σ_j is a function of downwind distance i and atmospheric stability parameter: $\sigma_j = \beta j/(1+\gamma j)^{-1/2}$, where α is a coefficient that converts the computed XCO_2 enhancement in the unit of ppm, and β and γ are coefficients depending on the atmospheric Pasquill stability category which is a function of the wind speed and solar radiation (Turner, 1970). The values for β and γ can be found in Bowers et al. (1980). The original Gaussian plume model generates a stationary plume… "

2.3 Clumps
I don't like the terminology "emission clumps". It doesn't fit with the actual definition of a clump:
noun: "a compacted mass or lump of something"
verb: "form into a clump or mass"
Emissions don't clump. The various sources have just been grouped together. The abstract of

their 2019 paper seemed to use "hotspot" and "clusters" which I would prefer to "clump". A cluster would be a much more intuitive name for this.

Response:

In our 2019 paper (Wang et al., 2019), we used the word "emission clump", which was defined as "clusters of emitting pixels (called emission clumps hereafter) that will generate individual XCO_2 plumes that are detectable from space". Since we strongly link our paper to Wang et al. (2019), we believe, for clarity and consistency, that keeping the term "clump" is critical.

We can also mention that in Merriam-Webster's Collegiate Dictionary, one of the definition given for "clump" is "a group of things clustered together" (https://www.merriam-webster.com/dictionary/clump). So we think "clump" is still appropriate, in the context of American English.

2.4 References

The authors show a very strong bias towards European studies. They don't seem to mention any of Ray Nasser's work in the intro even though his 2017 GRL paper used a Gaussian plume model with satellite observations to study individual sources. They also seem to have missed Eric Kort's work using GOSAT to study megacities (Kort et al., GRL 2012; among others).

Response:

We thank the reviewer to remind these references. In the revised introduction, we rewrite the paragraph setting the context for XCO_2 plume inversions:

Ln 55 "… Alternatively, vertically integrated columns of dry-air mole fractions of CO_2 (XCO_2) from satellites offer the opportunity to sample the atmosphere with a global coverage. Kort et al. (2012) and Janardanan (2016) found that significant XCO_2 enhancements could be detected over some megacities using Greenhouse Gases Observing Satellite (GOSAT) XCO_2 observations. Schwandner et al. (2017) also found XCO_2 enhancements of 4.4 to 6.1 ppm in the Los Angeles urban CO_2 dome using observations from Orbiting Carbon Observatory-2 (OCO-2). Nassar et al. (2017) used the XCO_2 observations from OCO-2 to quantify CO_2 emissions from several middle- to large-sized coal power plants. However, the design of GOSAT and OCO-2 observations with sparse sampling was mainly focused on the monitoring of CO_2 natural fluxes. Recent studies show a limited amount of clear detections of transects of XCO_2 plumes from cities or plants in OCO-2 observations (Zheng et al., 2020) so that GOSAT and OCO-2 data keep on being hardly used to estimate CO_2 city emissions. The potential for reducing uncertainties in fossil fuel CO_2 emissions at the scale of point sources (Bovensmann et al., 2010), cities (Broquet et al., 2018; Pillai et al., 2016) and agglomerations of several cities (O'Brien et al., 2016) should dramatically change with the planned satellite missions with imaging capabilities. These studies consistently showed that …"

2.5 3 hours vs 6 hours

Why is there a 6-hour window for Paris and a 3-hour window globally? I see, it's defined afterward. This should be moved forward to explain why Broquet chose 6 hours and why they choose 3 hours. How is 3 hours chosen? It seems to just be picked randomly.

Response:

Broquet et al. (2018) showed that the XCO_2 signature of the emissions from Paris is hardly detectable after 6 hours due to atmospheric diffusion, and they thus only inverted emissions during the 6 h before satellite overpasses. In PMIF-Paris experiments, we aim to compare the performance of inversion system using a Gaussian plume model with the one using a 3-D Eulerian atmospheric transport model, so we choose the same time length as Broquet et al. (2018) for PMIF-Paris. For PMIF-Globe, we already explained in the manuscript (in the revised version, we bring the explanation to Sect. 2.1, see below). On the other hand, three hours is the typical time scale that Nassar et al. (2017) used to interpret the results from their inversion of emissions from coal power plants using OCO-2 observations with a Gaussian plume model.

In the revised manuscript, we bring the explanation about the 6-hour time window for PMIF-Paris and 3-hour time window for PMIF-Globe to Sect. 2.1:

Ln 157: "Table 1 and 2 summarize the different options for the configuration of the system and of the OSSEs. One distinction between PMIF-Paris and PMIF-Globe is that PMIF-Paris relates XCO_2 signals with the mean emissions 6 hours before overpasses, while it is assumed that in PMIF-Globe that the XCO_2 signals only provide effective constraints on 3 h mean emissions before individual overpasses. The 6-hour period corresponds to the period of emissions from Paris whose signature in the XCO_2 field can still be detected by the satellite despite the atmospheric diffusion (Broquet et al., 2018). While Broquet et al. (2018) indicated that the period of "detectable" emissions from a large megacity like Paris could last up to 6-hours, most of the clumps across the globe have smaller emission rates than Paris, or are located in more complex environment close to other major emission areas where XCO_2 signals can be attributed to multiple sources, making the detection of the XCO_2 signature of emissions few hours before the satellite overpass even more difficult. For the PMIF-Globe experiments, we thus conservatively assume that the XCO_2 signals can only provide effective constraints on 3 h mean emissions before individual overpasses in general."

We also rewrote the paragraph in Sect. 2.3:

Ln 179-Ln186: "In the PMIF-Paris inversion, the satellite observations are sampled at 11:00 local time, in line with the experiments from Broquet et al. (2018). The inversion solves for the mean emissions for the 6 hours before 11:00 local time. Broquet et al. (2018) solved for the hourly emissions during this 6-hour period but PMIF can only solve for the mean emissions during the 6-hour period due to the fact that the Gaussian plume model cannot be used to compute the signatures in the XCO_2 field of individual hourly emissions during that period. The control parameter in PMIF-Paris for each overpass (Sect. 2.7.1) is thus a scaling factor λ for the mean emission between 05:00 and 11:00 …"

3 Specific comments:

Title: Remove fossil fuel from the title. I don't see how they could differentiate fossil from non-fossil sources in their analysis.

Response:

In this study, all the inversions and discussions focus on fossil fuel CO_2 emissions since this should be the main target of CO_2 emission monitoring systems, and since the PMIF is based on an inventory of these emissions and assumes that uncertainties in other fluxes weakly impact the inversion of these emissions in clumps. However, we agree that the separation between fossil fuel emissions and non-fossil CO_2 fluxes is a critical topic for the space-borne (and more

generally atmospheric) monitoring of the fossil fuel emissions. Firstly, background concentrations around the plumes from fossil fuel emission clumps might be sometimes difficult to properly separate (Kuhlmann et al., 2019). This background consists in a mix of the signature of all kind of CO_2 fluxes outside or within the clump boundaries. However, in a general way, uncertainties in this background can be seen as a source of uncertainty in the estimate of the fossil fuel emissions that does not prevent us from computing the fossil fuel emissions separately. Secondly, if focusing on sources and sinks collocated with the fossil fuel emissions for cities, the separation of fossil fuel emissions from biofuel emissions, human respiration and potentially natural fluxes specific to urban areas (i.e. highly different from natural fluxes at larger scale) can definitely be difficult. We investigated some estimates of the contribution of non-fossil CO_2 fluxes to the total CO_2 fluxes from cities. The contribution of non-fossil CO_2 fluxes to the total CO_2 fluxes varies a lot from city to city and from day to day. For example, in Île-de-France, the biogenic fluxes are usually considered to have small impact on the signals of fossil fuel CO_2 emissions in autumn and winter, while they could become non-negligible in summer (Bréon et al., 2015; Lian et al., 2019; Staufer et al., 2016); The biogenic CO_2 fluxes could represent 5% of the total signals in Indianapolis, Indiana, U.S.A. (Turnbull et al., 2015) during winter time; Miller et al. (2018) estimated that biogenic CO_2 fluxes could contribute to 25% of the total CO_2 enhancement in the Los Angeles Basin based on atmospheric radiocarbon measurements; Ye et al. (2020) estimated the contribution of total XCO_2 enhancement due to biogenic fluxes can be as large as $32 \pm 27\%$ (1σ) and $24 \pm 18\%$ (1σ) in winter and summer. All these estimates include the urban and rural areas, while the emission clumps defined in Wang et al. (2019) only include the areas with fossil fuel CO_2 emissions being high enough to form detectible XCO_2 plumes through atmospheric transport. Most of these areas are built-up areas, so the contribution of non-fossil CO_2 fluxes to the total fluxes should be much smaller than the whole-city estimates as mentioned above. This can be illustrated by Fig. 4a in Lian et al. (2019) of the small biogenic fluxes in the city center of Paris and by Fig. 1 in Ye et al. (2020) of the green vegetation fraction. We thus assume that in these clump areas, the fossil fuel CO_2 emissions dominate the total CO_2 fluxes.

In summary, we do agree with the reviewer that the satellite observations alone do not separate the fossil fuel emissions and non-fossil fuel fluxes within or around emission clumps and that these non-fossil fuel fluxes can be non-negligible. However, as shown by previous studies, the impact of non-fossil sources is within the overall uncertainty of the estimates of emissions from real data (Reuter et al., 2019; Zheng et al., 2020).

In the revised manuscript, we discussed the impact of non-fossil fluxes in more detail:

Ln 519-523: "…Broquet et al. (2018) showed that systematic error could hamper the ability of the inversion system to reduce the errors in the emissions estimates. Thirdly, we neglect the impact of uncertainties in diffuse fossil fuel CO_2 emissions (outside clumps) and non-fossil CO_2 fluxes (within and outside clumps), the latter including net ecosystem exchange (NEE) from the terrestrial biosphere, the CO_2 emitted by the burning of biofuel, the respiration from human and animals (Ciais et al., 2020) and the net CO_2 fluxes between the atmosphere and ocean. For example, the signals from terrestrial NEE can be strong during the growing season, and the signals from ocean CO_2 fluxes may have a critical impact on the overall XCO_2 patterns in the proximity of coastlines. In principle, the signals of diffuse fossil fuel CO_2 emissions and non-fossil CO_2 fluxes outside the clumps can be potentially filtered by removing

the local background XCO_2 field to extract plumes generated only by emissions from clumps (Kuhlmann et al., 2019; Reuter et al., 2019; Ye et al., 2020; Zheng et al., 2020). The non-fossil CO_2 fluxes within clumps vary from clump to clump, and could contribute a non-negligible fraction of the total CO_2 fluxes in many clumps (Bréon et al., 2015; Ciais et al., 2020; Wu et al., 2018a). The satellite observations alone cannot effectively differentiate the fossil fuel CO_2 emissions and the non-fossil CO_2 fluxes within clumps. In the clumps with non-negligible non-fossil CO_2 fluxes, the inversion of fossil fuel CO_2 emissions could be influenced (Ye et al., 2020; Yin et al., 2019). Fourthly, …"

Section 2.1: Should reference the sections that define the error covariance parameters.

Response:

We revised the manuscript:

- Ln 145: "We characterize **B**, **R** and **A** by the corresponding standard deviations (σ) of uncertainty in individual or aggregations of control parameters and by the temporal auto-correlations of the uncertainties (Sect. 2.6). In the following, …";

- Ln 154-157: "… Then we apply the system to all the emission clumps over the globe and over 1 year using a different control vector and a simulation of the XCO_2 sampling by a single CO2M satellite (Sect. 2.2). The inversions for all emission clumps over the globe are called PMIF-Globe. In PMIF-Globe, we first investigate the potential of satellite observations in constraining emissions from individual days (ExpNoCor in Sect. 2.6). Then we assess the ability of satellite observations to constrain emissions at annual scale by accounting for the temporal auto-correlation of the prior uncertainties (other experiments in Sect. 2.6). Table 1 and 2 summarize the different options for the configuration of the system and of the OSSEs."

Line 126: what is y_{fixed}?

Response:

We revised the sentence:

"… The inversion derives a statistical estimate for a set of control variables x in a model $x \rightarrow y = \mathbf{M}x$ that simulates the satellite XCO_2 measurements y^o. The model **M** linking x and y is a combination of flux and atmospheric transport models (detailed in Sect. 2.4), and is called observation operator hereafter. As explained below, we do not have a constant term added to **M**x in the observation operator of the PMIF that would gather the atmospheric CO_2 signature of the fluxes not controlled by the inversion (like non-fossil fluxes and the background XCO_2 field) since the uncertainty in such fluxes is ignored. The inversion follows a Bayesian statistical framework,…"

Line 181: rephrase, too colloquial: "but the PMIF can hardly handle hourly emissions when covering a whole year".

Response:

We revised the sentence:

"…Broquet et al. (2018) solved for the hourly emissions during this 6-hour period but PMIF can only solve for the mean emissions during the 6-hour period due to the fact that the Gaussian plume model cannot be used to compute the signatures in the XCO_2 field of individual

hourly emissions during that period. The control parameter for each overpass …"

References

Boesch, H., Baker, D., Connor, B., Crisp, D. and Miller, C.: Global Characterization of CO2 Column Retrievals from Shortwave-Infrared Satellite Observations of the Orbiting Carbon Observatory-2 Mission, Remote Sensing, 3(2), 270–304, doi:10.3390/rs3020270, 2011.

Bowers, J. F., Bjorklund, J. R., Cheney, C. S. and Schewe, G. J.: Industrial source complex (ISC) dispersion model user's guide, US Environmental Protection Agency, Office of Air Quality Planning and Standards., 1980.

Bréon, F. M., Broquet, G., Puygrenier, V., Chevallier, F., Xueref-Remy, I., Ramonet, M., Dieudonné, E., Lopez, M., Schmidt, M., Perrussel, O. and Ciais, P.: An attempt at estimating Paris area CO2 emissions from atmospheric concentration measurements, Atmospheric Chemistry and Physics, 15(4), 1707–1724, doi:10.5194/acp-15-1707-2015, 2015.

Broquet, G., Chevallier, F., Rayner, P., Aulagnier, C., Pison, I., Ramonet, M., Schmidt, M., Vermeulen, A. T. and Ciais, P.: A European summertime CO2 biogenic flux inversion at mesoscale from continuous in situ mixing ratio measurements, J. Geophys. Res., 116(D23), doi:10.1029/2011jd016202, 2011.

Broquet, G., Bréon, F.-M., Renault, E., Buchwitz, M., Reuter, M., Bovensmann, H., Chevallier, F., Wu, L. and Ciais, P.: The potential of satellite spectro-imagery for monitoring CO2 emissions from large cities, Atmos. Meas. Tech., 11(2), 681–708, doi:10.5194/amt-11-681-2018, 2018.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F. M., Chédin, A. and Ciais, P.: Inferring CO2 sources and sinks from satellite observations: Method and application to TOVS data, J. Geophys. Res., 110(D24), doi:10.1029/2005jd006390, 2005.

Ciais, P., Wang, Y., Andrew, R., Bréon, F.-M., Chevallier, F., Broquet, G., Nabuurs, G.-J., Peters, G., McGrath, M., Meng, W., Zheng, B. and Tao, S.: Biofuel burning and human respiration bias on satellite estimates of fossil fuel CO2 emissions, Environ. Res. Lett., doi:10.1088/1748-9326/ab7835, 2020.

Cogan, A. J., Boesch, H., Parker, R. J., Feng, L., Palmer, P. I., Blavier, J.-F. L., Deutscher, N. M., Macatangay, R., Notholt, J., Roehl, C., Warneke, T. and Wunch, D.: Atmospheric carbon dioxide retrieved from the Greenhouse gases Observing SATellite (GOSAT): Comparison with ground-based TCCON observations and GEOS-Chem model calculations, Journal of Geophysical Research: Atmospheres, 117(D21), doi:10.1029/2012JD018087, 2012.

Janardanan, R., Maksyutov, S., Oda, T., Saito, M., Kaiser, J. W., Ganshin, A., Stohl, A., Matsunaga, T., Yoshida, Y. and Yokota, T.: Comparing GOSAT observations of localized CO2 enhancements by large emitters with inventory-based estimates, Geophys. Res. Lett., 43(7), 2016GL067843, doi:10.1002/2016GL067843, 2016.

Kort, E. A., Frankenberg, C., Miller, C. E. and Oda, T.: Space-based observations of megacity carbon dioxide, Geophys. Res. Lett., 39(17), L17806, doi:10.1029/2012GL052738, 2012.

Kuhlmann, G., Broquet, G., Marshall, J., Clément, V., Löscher, A., Meijer, Y. and Brunner, D.: Detectability of CO2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO2 Monitoring (CO2M) mission, Atmospheric Measurement Techniques Discussions, 1–35, doi:https://doi.org/10.5194/amt-2019-180, 2019.

Lian, J., Bréon, F.-M., Broquet, G., Zaccheo, T. S., Dobler, J., Ramonet, M., Staufer, J., Santaren, D., Xueref-Remy, I. and Ciais, P.: Analysis of temporal and spatial variability of atmospheric CO2 concentration within Paris from the GreenLITE^{TM} laser imaging experiment, Atmospheric Chemistry and Physics, 19(22), 13809–13825, doi:https://doi.org/10.5194/acp-19-13809-2019, 2019.

Miller, J. B., Lehman, S., Verhulst, K. R., Miller, C. E., Duren, R. M., Yadav, V., Newman, S. and Sloop, C.: Unexpected and significant biospheric CO2 fluxes in the Los Angeles Basin indicated by atmospheric radiocarbon (14CO2), in 45th Global Monitoring Annual Conference, vol. 2018, pp. A53F–03, Boulder, Colorado, USA., 2018.

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A. and Crisp, D.: Quantifying CO2 Emissions From Individual Power Plants From Space, Geophys. Res. Lett., 44, 10045–10053, doi:10.1002/2017GL074702, 2017.

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O. and Wunch, D.: The ACOS CO2 retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmos. Meas. Tech., 5(1), 99–121, doi:10.5194/amt-5-99-2012, 2012.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T. and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO2 inversions, Biogeosciences, 10, 6699–6720, doi:10.5194/bg-10-6699-2013, 2013.

Rayner, P. J., Utembe, S. R. and Crowell, S.: Constraining regional greenhouse gas emissions using geostationary concentration measurements: a theoretical study, Atmospheric Measurement Techniques, 7(10), 3285–3293, 2014.

Rayner, P. J., Michalak, A. M. and Chevallier, F.: Fundamentals of data assimilation applied to biogeochemistry, Atmospheric Chemistry and Physics, 19(22), 13911–13932, doi:https://doi.org/10.5194/acp-19-13911-2019, 2019.

Reuter, M., Buchwitz, M., Schneising, O., Krautwurst, S., O'Dell, C. W., Richter, A., Bovensmann, H. and Burrows, J. P.: Towards monitoring localized CO2 emissions from space: co-located regional CO2 and NO2 enhancements observed by the OCO-2 and S5P satellites, Atmospheric Chemistry and Physics, 19(14), 9371–9383, doi:https://doi.org/10.5194/acp-19-

9371-2019, 2019.

Schwandner, F. M., Gunson, M. R., Miller, C. E., Carn, S. A., Eldering, A., Krings, T., Verhulst, K. R., Schimel, D. S., Nguyen, H. M., Crisp, D., O'Dell, C. W., Osterman, G. B., Iraci, L. T. and Podolske, J. R.: Spaceborne detection of localized carbon dioxide sources, Science, 358(6360), eaam5782, doi:10.1126/science.aam5782, 2017.

Staufer, J., Broquet, G., Bréon, F.-M., Puygrenier, V., Chevallier, F., Xueref-Rémy, I., Dieudonné, E., Lopez, M., Schmidt, M., Ramonet, M., Perrussel, O., Lac, C., Wu, L. and Ciais, P.: The first 1-year-long estimate of the Paris region fossil fuel CO2 emissions based on atmospheric inversion, Atmos. Chem. Phys., 16(22), 14703–14726, doi:10.5194/acp-16-14703-2016, 2016.

Turnbull, J. C., Sweeney, C., Karion, A., Newberger, T., Lehman, S. J., Tans, P. P., Davis, K. J., Lauvaux, T., Miles, N. L., Richardson, S. J. and others: Toward quantification and source sector identification of fossil fuel CO2 emissions from an urban area: Results from the INFLUX experiment, J. Geophys. Res., 120(1), 292–312, 2015.

Turner, D. B.: Workbook Of Atmospheric Dispersion Estimeates. Office of Air Program Pub. No. AP-26, Environmental protection agency, USA., 1970.

Wang, Y., Broquet, G., Ciais, P., Chevallier, F., Vogel, F., Wu, L., Yin, Y., Wang, R. and Tao, S.: Potential of European 14CO2 observation network to estimate the fossil fuel CO2 emissions via atmospheric inversions, Atmos. Chem. Phys., 18(6), 4229–4250, doi:10.5194/acp-18-4229-2018, 2018.

Wang, Y., Ciais, P., Broquet, G., Bréon, F.-M., Oda, T., Lespinas, F., Meijer, Y., Loescher, A., Janssens-Maenhout, G., Zheng, B., Xu, H., Tao, S., Gurney, K. R., Roest, G., Santaren, D. and Su, Y.: A global map of emission clumps for future monitoring of fossil fuel CO2 emissions from space, Earth System Science Data, 11(2), 687–703, doi:https://doi.org/10.5194/essd-11-687-2019, 2019.

Wu, K., Lauvaux, T., Davis, K. J., Deng, A., Coto, I. L., Gurney, K. R. and Patarasuk, R.: Joint inverse estimation of fossil fuel and biogenic CO_2 fluxes in an urban environment: An observing system simulation experiment to assess the impact of multiple uncertainties, Elem Sci Anth, 6(1), 17, doi:10.1525/elementa.138, 2018a.

Wu, L., Broquet, G., Ciais, P., Bellassen, V., Vogel, F., Chevallier, F., Xueref-Remy, I. and Wang, Y.: What would dense atmospheric observation networks bring to the quantification of city CO2 emissions?, Atmos. Chem. Phys., 16(12), 7743–7771, doi:10.5194/acp-16-7743-2016, 2016.

Wu, L., Hasekamp, O., Hu, H., Landgraf, J., Butz, A., aan de Brugh, J., Aben, I., Pollard, D. F., Griffith, D. W. T., Feist, D. G., Koshelev, D., Hase, F., Toon, G. C., Ohyama, H., Morino, I., Notholt, J., Shiomi, K., Iraci, L., Schneider, M., Mazière, M. de, Sussmann, R., Kivi, R., Warneke, T., Goo, T.-Y. and Té, Y.: Carbon dioxide retrieval from OCO-2 satellite observations using the RemoTeC algorithm and validation with TCCON measurements, Atmospheric

Measurement Techniques, 11(5), 3111–3130, doi:https://doi.org/10.5194/amt-11-3111-2018, 2018b.

Ye, X., Lauvaux, T., Kort, E. A., Oda, T., Feng, S., Lin, J. C., Yang, E. G. and Wu, D.: Constraining Fossil Fuel CO2 Emissions From Urban Area Using OCO-2 Observations of Total Column CO2, Journal of Geophysical Research: Atmospheres, 125(8), e2019JD030528, doi:10.1029/2019JD030528, 2020.

Yin, Y., Bowman, K., Bloom, A. A. and Worden, J.: Detection of fossil fuel emission trends in the presence of natural carbon cycle variability, Environ. Res. Lett., 14(8), 084050, doi:10.1088/1748-9326/ab2dd7, 2019.

Yoshida, Y., Ota, Y., Eguchi, N., Kikuchi, N., Nobuta, K., Tran, H., Morino, I. and Yokota, T.: Retrieval algorithm for CO2 and CH4 column abundances from short-wavelength infrared spectral observations by the Greenhouse gases observing satellite, Atmos. Meas. Tech., 4(4), 717–734, doi:10.5194/amt-4-717-2011, 2011.

Zheng, B., Chevallier, F., Ciais, P., Broquet, G., Wang, Y., Lian, J. and Zhao, Y.: Observing carbon dioxide emissions over China's cities and industrial areas with the Orbiting Carbon Observatory-2, Atmospheric Chemistry and Physics, 20(14), 8501–8510, doi:https://doi.org/10.5194/acp-20-8501-2020, 2020.